DataLine

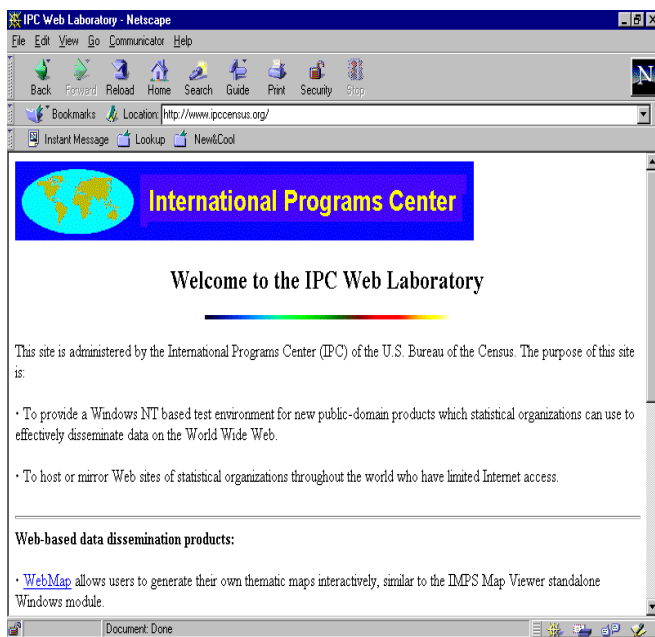## Another IPC on the Web

IPC has opened a second internet web site to showcase statistical information dissemination and some of our software.

http://www.ipccensus.org

Presently, the web site hosts **WEBMAP** and **WEBCROSS**, two products similar to Map Viewer and Cross Tabulation modules in IMPS.

In addition there are 'links' to other web sites which may be of some interest to our friends. If you have access to the WWW, please check it out and let us hear your comments.

In other news, our original web page has undergone some changes. Probably the most significant change is the ability to download maps of approximately 60 countries. Not only that, but you can get information on how to create 'map data files' which will allow you to create great looking thematic maps.

Basically, the idea is to designate a geographic area by a set of codes and then associate information (data) with the same set of codes. These two files are combined by Map Viewer to produce thematic maps. Just go to our web site:

http://www.census.gov/ipc/www/imps.html

and click on "Thematic Maps"; follow the instructions and you will be creating thematic maps in no time at all.

What's in a Name?

As you know, the IMPS and ISSA software teams have joined forces to create a new, more-powerful, easier-to-use statistical data processing product. The new 'processing system' will be available right after the millennium has debuted. In fact, the IMPS workshop scheduled for September 1999 at IPC will serve as a rigorous 'beta' test site for the newly-revised software. It will eventually combine the computational features of ISSA with the user-friendly features of IMPS. The first release will focus on a reengineering of the basic user interface. The Data Dictionary and Cross Tabulation modules, familiar to IMPS 4.1 users, will be restructured into an environment not unlike Windows Explorer. In addition, an initial version of the Data Entry module (the old CENTRY) will be available.

We are very excited about the new look of the combined product. However, we're in a quandry. . . What name shall we give to the new software? We have temporarily been calling it 'IMSA' for lack of a better name. We would like to come up with something completely different because the new system is NOT a combination of IMPS and ISSA but a new software package which combines the best characteristics of both while being significantly original.

What are your thoughts? What names can you suggest? We think the name should reflect the basic statistical nature of the software, but even that is open to challenge. Remember **CENTS** (**CEN**sus **T**abulation **S**ystem)? How about **CONCOR** (**CON**sistency and **COR**rection program)? Who could forget **TRS** (**T**able **R**etrieval **S**ystem)? In the same spirit, we are looking for a name that describes the function directly (like 'Statistical Processing Software') or represents the 'full' name through the use of initials (like '**I**ntegrated **M**icrocomputer **P**rocessing **S**ystem'). The name, of course, must also be unique, in the fine tradition of computer programming. We are fairly sure that we would run into trouble if we called it **PEPSI,** since that name is already spoken for.

Time is short, so if you have a name you think would be appropriate, either

1) send it via e-mail to imps@census.gov

or   2) fax it to IMPS at (301) 457-3033,

Attn: IMPS Design Staff

We really appreciate your help in finding a name. Watch the next DataLine for the result!

IPC Around the World

Ghana

Maria Mochulski, David Rain, Pat Anderson, Tom Ondra, Armando Levinson, and Nancy Murray worked closely with the Ghanaian Statistical Service (GSS) staff on their plans for the 2000 Population and Housing Census. **Stephen Adjei**, Director of the GSS Sampling Unit, and **Nana Akwasi Ango II,** also of the GSS Sampling Unit, are currently attending the IPC Sampling Workshop.

Tom conducted an IMPS workshop attended by **K.B. Danso-Manu, Seth Owusu Kwarteng, Abena Asamoabea Ani, Jacqueline Anum, Margaret Sallah-Richardson, George Wogblegbe**, and **Nobuyuki Ishii.** Ghana will be using IMPS to process its 2000 Population and Housing Census. During the workshop, Tom enjoyed sampling some Ghanaian delicacies such as fufu and kenke.

India

In February, Diana Lopez-Meisel worked with the Office of the Registrar General (ORGI) and met with the Director General, **Dr. M. Vijayanunni,** and many other acquaintances: **S.P. Sharma, R.G. Mitra, S.K. Sinha, A.K.Saxena, Minati Ghosh, Mahesh Ram, R.P. Singh, Himakar, K.N. Unni, T.S. Johar,** and **C. Chakavarty.**

In May, Selma Sawaya conducted a 4-week IMPS workshop at the Data Processing Division of ORGI. Participants included **B.L. Jain, Suraj Bhan, Raghuvansh Prasad, M.R. Balakrishnan, M.S. Thapa, Rajesh Kumar Dahiya, T.S. Johar, Anil Kumar, D.K. Sharma, Jagdish Chandra Joshi, Jaspal Singh Lamba, Ms. Usha**, and **Mrs. K. Malathi. Surendra Singh** provided invaluable assistance to his colleagues during the computer exercises.

Jordan

In April, a BUCEN team went to Jordan, where they met the new Director General of the Department of Statistics (DOS), **Dr. Hussein Shakhatreh.** The team worked with some of IPC's old friends **Fahad Hiyari, Fathi Nsour, Ikhlas Aranki, Khamis Raddad, Mohammed Al-Rifai, Zeinab El-Hasani, Jamal Jamil Sa'ad Al-Deen, Mohammed Al-Alami, Shamilla Murad, Abdel Wadoud Matouk, Mimer Hashem Nimer Gharbia, Wadji Akeel, Abed M. Awad**, and many others.

Malawi

IPC staff members Rebecca Sauer, Selma Sawaya, and Victoria Simmons visited the National Statistical Office (NSO) to finalize the data editing program and set up the control system and data entry operation to key the data from the 1998 Population and Housing Census. Sauer and Sawaya worked with **L. Mpando** and **L. Chipwatali** to finalize edit specifications; Simmons assisted **J. Thipa** and **A. Ndalira** to create the geographic database for the control system. **J. Kaphuka,** manager of the data entry operation, made sure that all the pieces would work together to produce quality information.

Mozambique

The operators finished keying the data from the 1997 Census of Population and Housing after 14 months of work. IPC long-term advisor,

Chris Corlett, and counterpart, **Tomas Bernardo,** are working on consistency and tabulations for the last three provinces of Nampula, Zambezia and Niassa. Final national results will be ready by the end of July.

Corlett and counterparts continue to produce electronic versions of the census results as each province becomes available. Ultimately, a CD-ROM containing the complete set of census publications will be available.

Philippines

One of the largest IMPS workshops ever was conducted at the NSO in Manila. **Tomas Africa**, NSO Administrator, **Nelia Marquez,** NSO Deputy Administrator, and **Klaus Beck**, UNFPA, joined the participants and many others in the festivities of the closing ceremony. The course was presented by Mike Stroot (IPC) and **Adrian Asilo** (NSO). **Ms. Au Reolalas**, UNFPA Project Coordinator, and her staff were extremely helpful in making sure that things moved smoothly. **Ms. Hilda (Datz) Ballesfin** and her group worked equally hard to ensure that everything required by the instructor or the participants was available in a timely fashion. The participants, all of whom worked very hard, included **Mae Almonte, Erma Aquino, Efren Armonia, Ysmael Baraguir, Glenn Barcenas, Magdalena Bautista, Stephen Bello, Maribel Bernardo, Norman Bundalian, Joseph Cajita, Reynario Capatoy, Adrian Cerezo, Joel Germono, Nicolas Gultiano, Jimmy Hinaut, William Jaro, Flordeliza Luistro, Cristina Mabanes, Elpidio Maramot, Sharon Martinez, Johnny Martinez Jr., Ray Merjilla, Gloria Miguel, Raymond Nimeno, Jeffrey Ocaya, Melchor Olmos, Heidy Palencia, Marifi Pedrera, Wilma Perante, Rona Perez, Melanie Rios, Edgar Siganay, Dennis Sorino, Gerardo Taguibolos, Amador Trazo, Naser Usman, Marilou Villanera.**

U.S. Virgin Islands

Dr. Frank Mills, Eastern Caribbean Center, welcomed Mike Levin and Mike Stroot to St. Thomas. They worked to put the finishing touches on a Household Income and Expenditure Survey. The trio was ably assisted by **Annette Gumbs,** who isolated certain problem areas so that they could be corrected.

Pacific Islands IMPS Activities

In December 1997, the United Nations Statistics Division funded a conference in Nadi, Fiji, to look at possible coordination of census questionnaire development and data processing for the 2000 Census round for the Pacific Islands. Representatives from 17 Pacific Islands nations and various international organizations attended the conference. **Mr. Laurie Lewis**, then of the UNFPA Country Support Team, was assigned the task of developing the core items, and **Carlos Ellis**, of the United Nations Statistics Division, and Michael Levin, of IPC, began work on proposals for common data processing requirements. The conference decided to use IMPS as the standard for the 2000 Census data processing. Ellis and Levin coordinated an IMPS training course at the University of the South Pacific in Suva in July 1998, attended by about 20 participants from more than 10 Pacific Islands countries and territories.

In May 1999, UNFPA funded a follow-up meeting in Suva, an "Expert Group Meeting on Model Population and Housing Questionnaires for the 2000 Round of Censuses in the Pacific." The following Pacific Islanders

IPC Around the World

attended: **Ms. Amelia Ngatokurua** (Cook Islands), **Ratu Timoci Bainimarama** (Fiji), **Kishor Chetty** (Fiji), **Ms Vasemaca Lewai** (Fiji), **Serevi Baledrokadroka** (Fiji), **Tekena Tiroa** (Kiribati), **Ms Lina Bade** (Papua New Guinea), **Magele Crawley** (Samoa)**, Ms Seini Filiai** (Tonga), **Taai Katarakei** (Tuvalu), and **Simil Johnson** (Vanuatu). Also attending were: **Laurie Lewis** (UNFPA Consultant), **Bill House** (UNFPA Country Support Team), **Iqbal Alam** (United Nations Statistics Division), **Dr Chris McMurray** (Secretariat of the Pacific Community), and **Jose Ferraris** (UNFPA Representative, Sub-Regional Office, Suva). Levin gave an IMPS demonstration during the meetings.

Two countries have already taken 2000 round censuses – Fiji and Tonga, both in 1996. Both used CONCOR for editing and CENTS for tabulation. All Pacific Islands nations expect to use IMPS for processing their censuses. Also, IMPS will be used in all of current and former U.S. Pacific Insular Areas for the 2000 round – the 1999 Marshall Islands Census, and the 2000 Censuses of Guam, Commonwealth of the Mariana Islands, American Samoa, and Palau. If the Federated States of Micronesia's census is funded, it will also use IMPS.

The U.S. Office of Insular Affairs continues to fund surveys and administrative records [which go into yearbooks] using IMPS processing. In March and April 1999, the following participants learned and used IMPS in Washington: **Gil Suguitan** (Guam), **Annette Gumbs** (Virgin Islands), **Tony Loa** (American Samoa), **Ayles Sukrad** (Palau), **Shannon Oseked** (Palau), **Jenkins Mariur** (Palau), and **Brihmer Johnson** (Federated States of Micronesia). **Will Maui** (CNMI) also was in Washington at that time to use IMPS on four censuses and five surveys to look at changes in alien labor force participation over time. During the last year IMPS was used to finish processing several of CNMI's surveys, the 1997 Economic Censuses for the Insular Areas, the 1997 Agriculture Censuses for these areas, the 1997 Virgin Islands Household Income and Expenditures Survey, and it is currently being used on the 1999 Agriculture Census for American Samoa.

The Pacific Islands nations and territories have committed to using IMPS for the 2000 Round of Census processing. Many of the countries are also beginning to use IMPS for various surveys and administrative record keying as well.

Visitors to IPC

Some visitors from the Bureau of Labor Statistics received an IMPS demonstration as part of a larger program. **Abdullah Al-Mishari,** Saudi Arabia, **Akram Mohamed**, Egypt, **Donald Prohaska**, Bosnia, **Muhammad Shafiqur Rahman**, Bangladesh, and **Janka Rovnakova**, Slovakia, were most impressed by the MapViewer which allows the presentation of data as thematic maps.

From Jordan, Department of Statistics (DOS) systems analyst and data base designer **Alaa' Ensheiwat** worked with Peter Johnson for nine weeks (April to June) on the development of a Windows-based interface for the Rural and Urban Projection System, a demographic analysis package.

For the Data Dissemination Workshop in May, from DOS we had **Jamil I.M. Hamdan;** from the Social Security Corporation we had two persons: **Nihaya Elayyan** and **Yahia Mustafa;** and from the Jordan National Population Commission, we had **Nayef Irteimeh.** For

Sampling and Statistical Methods Workshop in June, we had three Jordanians from DOS: **Abdel Wadoud Rebhi Matouk, Amer Al-Momany**, and **Hwida I.M. Al-Mansour.**

From India, for the Sampling and Statistical Methods Workshop, we had **S. Sanyal** and **Hari Krishan** from the Office of the Registrar General.

Monina Collado, Lourdes Hufana, and **Fiel Pizarro**, from the National Statistics Office of the Philippines, and **Flor Mateo** of the University of the Philippines, visited the Census Bureau as part of a study tour on data dissemination and social mobilization. They are preparing for their 2000 Census of Population and Housing. The group also visited several other institutions around the United States.

Y2K - Is It a Problem?

There has been much publicity about the so-called Year 2000 'bug' in computers. The basic problem is whether your computer will recognize the date 1/1/00 as the year 2000 or 1900. The latter causes all the dates stored after 1/1/2000 to be used incorrectly and, in some cases, these dates may even be treated as illegal or invalid.

Does your computer have this problem? Some people will wait until January 1, 2000, then turn on their computer and see if it works! If you would rather find out before then, you need to know how to do it. We have found four sites that allow you to download a free Y2K test:

> http://www.accute.com
>
> http://www.y2kbase.com
>
> http://www.hughesnet.net
>
> http://www.zdnet.com/vlabs/y2k/testy2k.html

∟Don't wait until the last minute of the millennium to do something!

Y to K Completed

Our crack staff have finished work on the Y-to-K conversion and, to be honest, we really do not see what the problem is! Changing 'Y' to 'K' is not that difficult.

For example: Here, for the first time, is a sample of the revised Y-to-K compliant calendar

K e a r   2 0 0 0
J a n u a r k

Sundak	Mondak	Tuesdak	Wednesdak	Thrusdak	Fridak	Saturdak
						1
2	3	4	5	6	7	8

The entire calendar is available for $10,000 U.S. per copy. (The price may seem high but we need to recover our costs.)

Reader's Corner

CENTRY Range Checks

Dear *DataLine*,

A colleague who is trying to build a data entry application submitted the following "problems" to me:

1- even though he defined ranges for common variables in the dictionary, any value is accepted while entering data for these variables.

2- "blank" value is accepted for numeric items in the dictionary but is not tolerated while entering data! (see Q7).

Attached are FS99.DD and FS99.AP. The explanation might be in the manuals but it's faster to ask!!! Thanks a lot,

Sabah El Mrini, Morocco

Dear Sabah,

The problem is that the changes to the dictionary must be incorporated in the AP. Go to CENTRY, then DEVELOP APPLICATION again with DD and AP. A red screen will come up with ?? Range changes. This means that the ranges in the AP have been changed to the ranges in the DD. Save the AP and all should be well.

This is a common problem which we will try to address in the Windows version. IN DOS, the DD and AP are separate files. The AP contains the DD which developed it. Changes to the DD DO NOT automatically change the AP, it must be re-developed. Sometimes this is good, sometimes NOT so good. As always, users should make back-ups of DD and AP before making changes.⌡

Inserting cards in CENTRY entry mode

Dear *DataLine*,

I attempted to add 2 records in my batch: c141 and c151 , I hit F3 then F5 to get a list to choose from, but c141 and c151 were skipped!! I tried the CLEANAP utility but that didn't help. Attached are menages.dd, menages.ap, and a sample file: test.bch. Thanks for answering promptly!!!

Sabah El Mrini

Dear Sabah,

This IS a problem. We have reproduced it and it has to do with some changes to the CENTRY which were made when the limit of record types was increased from 50 to 100. What can you do to 'fix' it? Well, this is the tough part. We have a new CENTRY EX.EXE which the IMPS system team thinks will solve your problem. We are sorry that it took this long but as we say "Better late than never". Attached is a new EX.EXE which will replace your 'old' CENTRY Executor. Make a copy of the 'old' EX.EXE so that you will have it as a backup then replace it in the \IMPS31\CE\ directory with the 'new' one. Please test it out and let us know. We have NOT fully tested it yet because we wanted to get you back to work. It appears to be OK but in the computer world you never know if it works completely right - you only know when it does NOT work! :

Another CENTRY problem

Dear *DataLine*,
Here am I again! A CENTRY problem was submitted to me and I could not figure out the solution.

1- the system displays the warning "updating the application to match the dictionary..." and then we would get a message about some "range changes" each time we opened the application even though there were no more changes in the dictionary.

2- for Z11, Z19 and Z21 fields (at least), the application does not behave properly in terms of accepting good values and rejecting bad ones. (see list of values in dictionary).

My guess is that there are too many values listed, and perhaps defining intervals would solve the problem? Attached are .DD and .AP files.

Sabah El Mrini

Dear Sabah,

You certainly do your best to keep us busy! This is the famous CENTRY 'range' problem, for which we have no real solution. We tried 'CLEANAP' on your application and got down to '1 range error'. You are correct about the number of values having something to do with it, so we 'compressed' some of the values from individual values to a range(example: 7011, 7012, 7013, & 7014 ---> 7011:7014) in the data dictionary.

After a couple of false tries, we guess we finally got under some mystical limit which seems to correct the problem. We have attached the 'new' DD and AP but given them new names. If you have more problems, try 'compressing' more of the values from individual codes to ranges-then run CLEANAP [in \imps41\imps31\ce folder], then re-develop the AP.⌡

Executing system developed in CONCOR

Dear *DataLine*,

Suppose we have completed our CONCOR validation program and we want to let a user execute it. How do we make it executable without passing through the IMPS menu? Will this be possible as it is not stated anywhere in the manual how to deploy the system to user?

For your information this validation module should be run immediately after the entry module (developed in Visual Basic). One of the buttons is to validate the data. It is done when the data capture is complete. This button is linked to IMPS. If possible, we do not want users to see the IMPS menu. How do we do it?

Laila Hafizan, Malaysia

Dear Laila,

This is a very good question. Since CONCOR is a DOS-based system, it is activated though a set of 'BAT's. You need to create your own BATs to use the CONCOR BATs.

We assume you are running Windows 95 and we assume that the data files will have a common extension -e.g. .DAT / .DTA, and the edited file (if you are using one) will have a fixed extension. e.g. OUT/ EDT. There is usually some problem with environment space so we have found it best to create two BATs:

 # 1: ---> RUN1.BAT [of course, you can pick any name for the BAT]
ECHO OFF

```
REM *** FILENAME: ?????.BAT
REM *** DESCRIPTION: Call BAT to run CONCOR validation
REM *** AUTHOR: ?????????  <------put name here!
REM *** DATE: 5/99
cd \WORK
command  /e:4096  /c  run2.bat  %1
```

NOTE: \WORK is directory where data files are located.
Run2.bat will require full path if not in \WORK.
%1 allows for parameter to be passed (name of data file).

#2: ----> RUN2.BAT

```
ECHO OFF
REM *** FILENAME: ????2.BAT
REM *** DESCRIPTION: Run CONCOR validation for census
REM *** INPUT: %1.DAT (unedited data file - common extension)
REM *** OUTPUT: %1.OUT (edited data file -if you are using)
REM *** AUTHOR: ?????????
REM *** DATE: 5/99
path>holdpath.bat
path=\IMPS41\IMPS31;\IMPS41\IMPS31\CN;\WINDOWS\COMMAND
set impsnet=XYZ
call run edit-program  %1.DAT  %1.OUT NUL  %1  NUL
holdpath
```

NOTE: edit-program is your edit EXE. If not in \WORK then you must give the full path.This assumes that each machine has IMPS installed as usual. [If space is a problem, all you really need is the IMPS31 and IMPS31\CN directories/folders.]

This is the THEORY but we have used it many times in the past. As usual it needs to be tested until it gives the desired results.

In Windows 95, you can set up a 'short cut' using:
command line: c:\work\run1.bat %1   <--- or wherever BAT is stored
working directory: c:\work
after setting it up: use 'Properties' to give a data file prompt and use 'Close on exit' check box. Other options - full screen, etc.

We are sure you will have other questions once you try to set this up. We have someone who works with Visual Basic so we may even be able to help with those problems.

Data Dictionary Conversion & ID Check File

Dear *DataLine*,

Problem #1 - I am working with both IMPS3.1(DOS version) and IMPS4.1(Windows version). I use the Data Dictionary in IMPS 4.1 to enter labels and values into the data dictionary. The problem came up when I switched to the DOS version of the Data Dictionary. I find that I get a lot of 'duplicate values' messages preventing me from exiting the values screen. The only way to get out is to edit all the duplicates and then not save the changes. I need to retain the text of the labels which I entered in the Windows version. Another problem concerning values is how to use the complete value text in Windows version for reports and table production seeing that most of my value text have been truncated in DOS version.

Problem #2 - I need to set up an ID check file for my data entry application. The structure is as follows with number of digits in bracket.
Batch id: Prov (2)   ward (2)   ea (2)

Questionnaire id: Household (3)

The ID check file will be too big because of the number of households in each ea.   Is there an alternative?

Problem #3 - Are there any good reference manuals available on CD?

Lorraine Taboru, Solomon Islands

Dear Lorraine,

Problem #1 - solution - Do all modifications to the data dictionary in 4.1 (DDW). After changes are complete "Save as" a 3.1 'DD'. Never make changes to DD directly and never 'Open' the DD (DOS version) in IMPS 4.1 and then 'Save' as 4.1 (DDW).   This causes the 'truncated' value names to replace your long names plus it replaces your item labels with the item name.

The reason we allow users to 'open' a DD and then save it as a DDW is to let users convert their 'old' IMPS 3.1 DDs to 4.1 DDWs when they use it the first time. After, ALL changes should be done in 4.1.

If you want to see what is 'in' the DD then use 'List data dictionary' in the 3.1 Data dictionary menu. It is true that the value names will be truncated in IMPS 3.1 but they are only used with QUICKTAB. All the modules in 3.1 will work with the converted DD even though there are 'duplicate' value names.

In CENTS, 'Format' the table using exactly the text you want. Then you must develop a TAB program that puts the numbers into the table. The value names in the DD do NOT come into play.

Problem #2 - solution - Since you will be using 10 stand-alone machines "one for each province", if you want to use ID check file we suggest you use 10 files, one for each province. Each would contain the valid IDs for that province. Do NOT worry about 'too big', these files are actually fairly small compared to others.

Will you know all your questionnaire IDs? If not, you could make a good guess at the IDs.

ID check file would contain records coded as follows:
ppwweahhh _____
where pp= this province code, always the same for the file
*   ww = 01 to the number of wards for this province*
*   ea = 01 to the number of EAs for this ward*
*   hhh = 001 - max hhh value*

There are several ways to generate these files if you need to. Probably it would be best to generate all the combinations, then delete the ones you don't need. We can guide you if needed. It is important to remember that CENTRY will not allow a keyer to enter a questionnaire if the ID is not in the ID check file OR if the ID has already been keyed. If you have Questionnaire IDs in the ID check file which do NOT have questionnaires associated with them, it is somewhat of a pain but it can be handled.

Another important idea: if the ID keyed is rejected by the ID check file have the keyer make sure it was entered correctly but do not allow them to change the values until they find something that is accepted. You need a procedure to follow when ID is rejected. By the way, you can make a DDW and CENTRY AP to review the ID check file. This sometimes helps rather than just looking at the file

using a text viewer. Put the ID fields as common items and the date as an item on a record. The file is a flat ASCII file - it is NOT indexed.

Another approach: Do not use an ID check file but key in a batch and sequence number along with the questionnaire. Batch = 045 Sequence = 10 for example. After the province is completely keyed, use CONCOR to WRITE the questionnaire ID, the batch and sequence into a WRITE file. Make a small data dictionary for this file (common= ID, batch, & seq on the record) then 'sort' in IMPS 4.1 by ID. The sort will list the duplicated IDs in a report.

With this report, you can find the ID in the WRITE file to determine batch & sequences. Of course, the difficulty here is that you must wait for all the batches of a province to be entered. We can explain this further if you are interested. It may be something you want to do even though you use an ID check file.

Problem #3 - solution - We assume that you mean IMPS reference manuals. All we have are the user's guides for IMPS 3.1 and 'Getting Started' for IMPS 4.1. Both of these are available from the WWW.

http://www.census.gov/ipc/www/imps.html

We have the IMPS 3.1 manuals (Word Perfect 5.1) format available on diskette (same manuals as on the WWW). We have no other documentation. The class notes are available but they are pretty much ideas/concepts to be covered in the class and exercises. We can send them along if you wish.⌡

CONCOR Program Run Problem

Dear *DataLine*,

I was able to compile the REFORM.CN program without any problems. However, when I ran the program, I got the following error message:
REFORM.EXE - NOT FOUND
REFORM.ARF - NOT FOUND

What does this mean?

Tanya Jones, Bermuda

Dear Tanya
These message simply mean that the program was NOT compiled on the PC on which it was running. When you compile you create

- REFORM.EXE
- CONCOR.LST
- REFORM.MSG
-REFORM.ARF (optional)

You need to transfer at least these files (except CONCOR.LST) to the computer on which you will execute the program. It is usually best to actually compile the program on the PC which you are going to use.⌡

IMPS - CONCOR & COMMON items

Dear *DataLine*,

I will try to explain to you exactly what we want to do with CONCOR at the INE Portugal. I'm using IMPS to validate a census questionnaire with CONCOR.

I have a problem about the Common data items. My questionnaire has the structure below:

1 - One Building record - Identified with a geographic code (5 items) and a building code.

2 - A group of household records per building - Identified with a geographic code, building code and a household code.

3 - A group of family records per household - Identified with a geographic code, building code, household code and a family code.

4 - A group of person records per family - Identified with a geographic code, building code, household code, family code and a person code.

Must I define the geographic codes and building code as common data items or must I define all codes (geographic, building, household, family and person codes) as common data items? I have to verify the existence and consistency of various records. If I use the first option, can I get cumulative values per household and family? Thank you!

Carlos Valente, Portugal

Dear Carlos,

We think we have a scheme which will allow you to process the census with minimal programming. If you do not like this technique we have some others ideas we can suggest but they are more complicated. First, let us say that even after your lengthy explanation it is difficult to figure out everything you are trying to do. We understand some of the strategies but not all.

OK, the basic processing, as we understand it, has Buildings with a building code; Houses with building and house codes; families with building, house and family codes; and persons with building, house, family and person codes. These are four separate files and, as such, are very, very likely to have problems matching all the linked records together. So your basic problem is to make sure you do not have missing or duplicate records. This is a headache of the first magnitude!

Here is what we propose: Every record in CONCOR must have a set of ID codes on it. In this case that would be a Building code, House code, family code (and person code although this is NOT part of the Questionnaire ID).

CODES ------>>	Build	House	Family	Person
RECORDS				
Building	*nnnn*	*999*	*99*	*99*
House	*nnnn*	*nnn*	*99*	*99*
Family	*nnnn*	*nnn*	*nn*	*99*
Person	*nnnn*	*nnn*	*nn*	*nn*

where nnn and nn are codes not equal to 999 or 99

Assuming that all records are exported to IMPS with these special codes for the corresponding items, the four files can be sorted together using ASCENDING values. So, the result file should have persons first, followed by family, followed by house, followed by building records.

In the Data Dictionary, use '2' for Max records of BUILDING, HOUSE and FAMILY records. [This is to check for duplicate codes of the same record type. IF you know this is NOT a problem, then you can use '1' instead.] Try to make Max records for persons large enough to handle the families with the maximum number of persons.

In CONCOR, use Build item + house item + family item as QUESTIONNAIRE ID.
What happens during processing:

a) CONCOR reads and processes a family in house. You will probably want to increment counters for # of families, # persons, etc. You can check TYPE-COUNT (Family) <> 1 for missing or duplicate Family records. You can do whatever 'family' checks might be required.

Note: there will NOT be a House or Building record in the memory. You must 'save' any information from these families that you wish to use in variables.

b) After all the families (if any) in the house are processed, the House record should come in alone (or by 2 if it has duplicated codes). Any counters from families, etc. would be tested at this time. IF the specifications call for changing the values on the House record if they do not match, then you can do it here. [Of course, you cannot change values on the family records because they are not in the memory]. After the testing, you would reset the family-type counters to zero, etc. You would also increment counters for houses and save any other information for checking on the Building record.

c) If there are more 'Houses' in the Building, steps a) and b) would be repeated for each.

d) After all the Houses in Building have been read and processed, the Building record should be in the memory. (Again you can check for duplicates if you wish.) Here you do any checking which involves the inter-record checks with other record-types. You must have saved the information in variables, of course. You would also reset counters used to verify totals in Building record. In terms of the CONCOR routines:

For-new-quest
----- executed for a) each family, b) each house and d) each building since these are considered different questionnaires.
For-each(Family)
----- executed for a) each family record only
For-each(Person)
----- executed for a) each person record in the family
For-each(House)
----- executed for b) each house record only
For-each(Building)
----- executed for d) each building record only
At-end-quest
same as For-new-quest.

Again, Person and Family records will be in the memory at the same time but house records and building records will be processed by themselves. Does this sound like a possible strategy? We are not really sure about all the 'edits' you are trying to do but we think this should handle all of them as long as you are modifying data on the 'higher'-level records. ♩
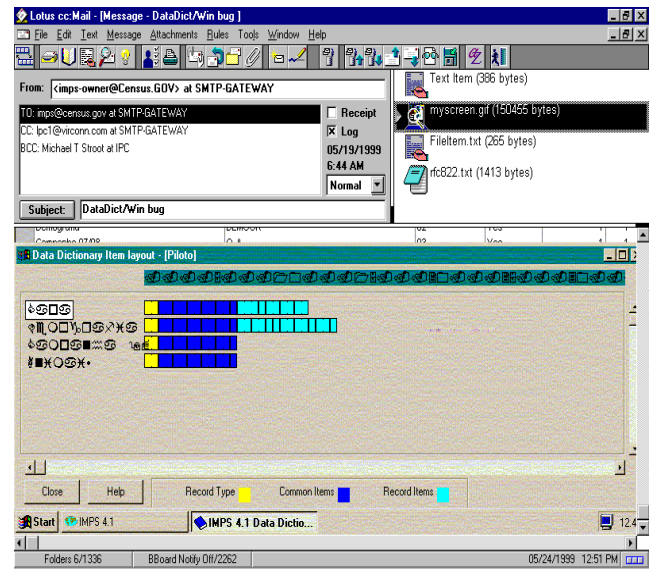
Data Dictionary/Win bug

Dear *DataLine*,

Have you ever seen the following bug in Windows Data Dictionary?

A user here is running IMPS 4.1 (December 98) under Win NT workstation (SP 3). When he goes into Layout Mode, the fonts for the Record Types and ruler show up in Wingding font. It's really weird!

Attached is a screen shot.

Chris Corlett, Mozambique

Dear Chris,

Yes, we've seen it. Yes, we've fixed it. The May 12 version of IMPS 4.1, which is on the Internet, has the fix. The bug was introduced when we added Armenian to the list of languages handled. The font change to MS Line Draw in the layout screen was slightly incorrect. Some operating environments of Win 95, Win98 or WinNT got mixed up and substituted Wingdings for MS Line Draw characters. But it is nice to know that someone noticed! ♩

Questions about IMPS

Dear *DataLine*,

We are processing a survey with IMPS and need your help in order to answer some questions that we cannot answer by ourselves. When we run CONCOR programs we get always the following message:

'This questionnaire was processed in parts at line 0 in questionnaire xxxx for each x'

We have read the Run Time Error Messages Appendix in the manual and have found:

'There is not sufficient room, MAX records, to store the entire questionnaire' and

'...increase MAX-RECORDS as needed so the entire questionnaire can be stored...'

We have done that in the CONCOR program but we are still getting the same warning. Do you have an explanation for it?

We have 16 types of records, of which 5 have Max-records defined greater than 1: we have one with 2 occurrences, two with 15 occurrences, one with 5 occurrences and one with 4 occurrences. The longest record is 137 characters and the shortest one is 41.

Furthermore, when we run QUICKTAB to produce tables, we get results, which are based on the whole sample if the data items we are crossing are in the same record type, but not otherwise. Can we do something different than to write all the data items we want to cross tabulate in the same record type, in order to get the right tables?

Alberto Camardiel, Venezuela

P.S. By the way, IMPS is a great software!!!

Dear Alberto Camardiel,

Thanks for the compliment about IMPS. We think it is the best software for processing surveys and censuses.

Now to your problems:

1) Questionnaire processed in parts.

This means exactly what you have said - there are not enough 'slots' in memory to store all the records of a certain type. Our guess is that it is one of those with MAX-RECORDS = 1. If a questionnaire has 2 of these records, they can be processed only one at a time.

How to FIND THE PROBLEM:

What we would do is write a simple CONCOR program to 'see' the offending questionnaire.

Control   Questionnaire xxxxx xxx etc.
Reports   Summary Quest

For-new-quest
If FULL-QUEST-FLAG <> "Y"
*   message "This questionnaire is in parts"*
*   DUMPQ*
End-if

What will happen is that this message will be executed twice (at least):
first - when the first 'half' of the questionnaire is stored in memory;
second - when the second 'half' of the questionnaire is processed.

In the QUEST REPORT, just look at the last record type in the first 'half' (and the first record type in the second half) and you should see which record type is causing the problem.
Note: this assumes that the record types occur in the same order in the data file as they are defined in the data dictionary.

2) QUICKTAB processing items on different records

This is more difficult. There is BAD news and GOOD news.

BAD news: QUICKTAB (IMPS 3.1) CANNOT cross items on different records.
GOOD news: Cross Tab (IMPS 4.1) can.

IMPS 4.1 is Windows 95- based software. Assuming you have Windows 95 and access to INTERNET:
a) go to our website:

http://www.census.gov/ipc/www/imps.html

b) download IMPS 4.1 or just Cross Tab

c) install it (IMPS 4.1 should find your Realia COBOL folder/directory)

d) open and save your IMPS 3.1 data dictionary as an IMPS 4.1 version.

e) run your Cross Tabs.

It is probably NOT that simple but we are here to help if you have any more problems.⏌

A Complete Processing System

Dear *DataLine*,

It is clear that I need some help in putting together my complete processing system. I want to understand what I am doing and test every step. Can you help me with this?

Peta Pasa, Western Samoa

Dear Peta,

We will suggest a system of processing but you can modify it as you see fit. We don't know your circumstances but this is a basic system which has worked well in other places.

1. Determine a method for naming batch files. This should be 1-8 characters from which you can tell where the batch is from and what it contains. Could be the geography codes plus month & day or something like that. Everyone must understand the coding scheme.

2. Set up bat's which keyers can use to enter & validate data. They don't need to go into IMPS to enter data, etc. All they need is the EX.EXE file from the \IMPS41\IMPS31\CE (CENTRY) directory. Make a directory called \ENTRY (or any other name you want) on each machine. Copy the EX.EXE and your CENTRY AP into it.

Use DOS edit (or any text editor) and create a bat, GO.BAT for example. In GO.BAT type the three lines

CD \ENTRY
IMPSNET iii
EX your.AP %1 NUL NUL NUL

IMPSNET iii are 3 initials which can be used on a network computer.
EX is the CENTRY executor.
*   your.AP is the name of your CENTRY AP.*
*   %1 allows for a variable Batch name.*

To use the bat, keyers would key in
*   GO batch-name   (Without BCH extension)*

If you have Windows 95/98 you can set up a shortcut and icon on the desktop to run this bat.

On the command line use:
*   \ENTRY\GO ?*

In this case the keyers could just double click on the icon - get a Windows box to enter the batch ID and off they would go.

The BCH would be created on their hard disk in the \ENTRY folder.

When the batch is complete they could copy it to a network directory either directly or use a BAT or use a shortcut.

    COPY %1.BCH G:\DATA\%1.BCH

assuming that G is a network drive and \DATA is a directory where you want to store the data. You could also use A: instead of G:\DATA to copy it onto a diskette.

We think it is important to make the keyer's job as simple as possible - when they are happy you get better data. After they have used the system a while, ask them if they are having any problems. Sometimes you can 'fix' the problems, but sometimes you can't. If they see you are trying to help, they will be encouraged to do a better job.

3. CENTRY and CONCOR

You are right about CONCOR not executing until the ENTIRE questionnaire has been entered. This was done for several reasons but the bottom line is - that's the way it works. How to turn this to your advantage? It's an entry-validation system that saves time and nerves.

A. Keyers enter data without CONCOR, the entire batch

B. Optional verification (you do not have to verify all fields nor all questionnaires). Each field can be verified or not (when you develop your AP put the cursor in the item field and see that F9 can be used as a switch to turn verification ON/OFF (YES/NO)) The default is YES.

C. A reviewer uses the BCH and the questionnaires and CONCOR program to make manual corrections.

- in CENTRY go to Modify Batch
- press F8 (CONCOR program will execute until a problem is found)
- the reviewer can make the corrections if necessary (F7 to accept changes) - press F8 again to find next problem. (If problem is NOT corrected, page down to next questionnaire BEFORE pressing F8 again.) The advantage here is that you don't slow down the data entry people - they don't run CONCOR.

Problems can be corrected by a person with some understanding of the survey. The reviewer can see what types of problems are occurring. Are the data entry people making a lot of mistakes, are the coders making mistakes, are the respondents making mistakes? Any recurring problems can then be addressed to the people making the mistakes. Again, what you are trying to do is get the information off the questionnaire and into the computer as accurately as possible. This is just another form of quality control, like verification.

So you might have BATs/shortcuts for
a) entering data using CENTRY
b) copying it to a network drive/diskette
c) copying it from a network drive/diskette for validation
d) validating data using CENTRY and CONCOR (BAT would have: EX your.AP %1 NUL your-concor.exe NUL)
e) copying validated batch to other diskette/network drive (G:\RDATA)   (this would keep the original keyed data as backup)

It would be good if you could tell from the file name if it is the original keyed data or the manually reviewed data. This is usually done by adding a letter to the name, but that means your original name can be at most 7 characters.

Example   0102FEB   would become R0102FEB - R means reviewed.

If you use a BAT this would be done by
    COPY %1.BCH G:\RDATA\R%1.BCH   (Or use A' for diskette)

This CONCOR program would be the one just looking for problems which would be manually corrected from the questionnaire. You have to consider what you want the reviewer to do if some essential data are missing or blank. They could make a 'guess,' or you could leave it blank and impute it in the next step.

4. Concatenating data files

After the data file has been reviewed you can use the concatenate files utility in IMPS. How do you want to concatenate?
If you just want to put them all together, then use the wild card option; but again, use a name for the concatenated file which tells you what it is - maybe something like month and year or whatever. For wild card, use R*.BCH and all the BCH files beginning with R will be included. Be very careful here that 1) you get all the files you want, and 2) you don't get extra files.

5. Automatic CONCOR

Once the BCH files are concatenated, run the final CONCOR program which fixes any problems and adds the recodes which you need for tabulations. The output of this program should be ready for Cross Tab or CENTS. Look carefully at the CONCOR reports to see if anything is drastically wrong. You should know how many imputations were made and what values were imputed.

We will help with this. If your entry and review steps are good, there should not be many problems. As the people learn about their mistakes, the imputations should almost vanish!

6. Cross Tab or CENTS

You can save the 'spec' file in Cross Tab so that all you have to do is call it up and run your tables. You can also write CENTS programs and run those if the tables are more complicated.

Of course, Cross Tab is much preferred if it meets your needs. Again we can help with this when the time comes.

This has been a long note but we think it is essential that you decide on the big picture before we go too much further on details. Once we decide on what we are doing, it will be easier to work on the pieces.

Other things to think about:

- what recodes do we want to put in the data dictionary?

- what checks do you want the reviewers to make (this would be the CONCOR program that goes with CENTRY)? Do you want them to check everything or just the most important things?

- How are you going to automatically fix problems with the data in the 2nd CONCOR program [Hot deck, not stated code, or use someone else in the household]? We think a good subject-matter person could help you if you need it. All this should give you SOMETHING TO DO TODAY!

Chile

El pasado 8 de marzo en una ceremonia oficial, se firmó un Acuerdo de Cooperación en Materias Estadísticas entre el Instituto Nacional de Estadísticas (INE) de la República de Chile y la Oficina del Censo del Departamento de Comercio de los Estados Unidos de América. El acuerdo fue firmado por el Director Nacional del INE, Sr. Máximo Aguilera Reyes y el Director de la Oficina del Censo, Sr. Kenneth Prewitt.

El propósito de este acuerdo es promover la colaboración y el intercambio en el campo de las estadísticas económicas, sociales, y demográficas en los países de habla hispana y portuguesa. El INE servirá como punto focal de irradiación coordinando y capacitando al personal de institutos estadisticos similares.

Atelier Régional de Formation en Méthodologie et Logiciels (IMPS) de Traitement de Données de Recensements Généraux de la Population et de l'Habitation (RGPH)

Le Bureau de Recensement des Etats Unis *(International Program Center)* et l'Equipe d'Appui Technique du FNUAP de Dakar organisent, avec l'appui du Bureau Central de Recensement du Sénégal, un Atelier Régional de Formation en Méthodologie et Logiciels (IMPS) de Traitement de Données de Recensements Généraux de la Population et de l'Habitation (RGPH), au Sénégal (Saly-Mbour) du 05 septembre au 02 octobre 1999.

Cet Atelier Régional de Formation s'inscrit dans le programme de L'EAT pour 1999 comme l'une des activités prioritaires, conformément aux recommandations de la rencontre de janvier 1999 entre les Représentants du FNUAP et les Conseillers Régionaux et conformément aux recommandations du PA/CIPD pour le renforcement des capacités nationales d'exécution des programmes de collecte de traitement, d'analyse, de diffusion et d'utilisation des données sur la population et le développement.

Cette activité est organisée pour pallier l'absence de cycle de formation en français dans le domaine du traitement informatique des données de RGPH ainsi que pour réduire le nombre et la durée des missions du conseiller en traitement de données qui sera particulièrement sollicité au cours des prochaines années dans le cadre des nombreux recensements de la série 2000 qui se réalisent dans la sous-région.

La formation des cadres nationaux au niveau régional permettra de réduire considérablement le coût comparativement à ce qu'il serait aux Etats Unis ou si la formation était répétée par le Conseiller régional au niveau de chaque pays.

L'atelier prendra en compte toutes les nouvelles recommandations pour la réduction des coûts et des délais de traitement des RGPH, pour une meilleure analyse / dissémination /utilisation des résultats ainsi que pour l'amélioration de la synergie entre les activités de collecte, de traitement informatique et d'analyse.

Nous invitons les bailleurs à prendre en charge la participation de deux ou trois cadres nationaux qui sont ou seront responsables du traitement informatique des données du Recensement Général de la Population et de l'Habitation (RGPH) en cours ou en préparation.

La formation, dispensée par Mme Selma Sawaya du Bureau de Recensement des Etats Unis et Robert De Clercq, Conseiller régional de l'EAT, assistés par deux cadres du BCR du Sénégal et M. André Mayouya, Conseiller régional de l'EAT, portera sur l'ensemble de la méthodologie du traitement informatique et des modules de la dernière version du logiciel IMPS (utilisé pour la plupart des RGPH), incluant les programmes et procédures de saisie (CENTRY), de contrôle de cohérence, de corrections automatiques (CONCOR) et de tabulation (CENTS et CROSSTAB).

Compte tenu de la technicité et de la complexité du programme, cet atelier s'adresse uniquement aux cadres nationaux ayant de solides bases en informatique et en programmation. La participation de cadres féminins et celle de statisticiens/démographes, répondant scrupuleusement à ces critères, est encouragée. L'atelier vise en priorité les pays qui n'ont pas pu participer à l'atelier IMPS de Bamako (1997) et ceux qui ont souffert, depuis, d'une mobilité des cadres.

Le caractère exceptionnel de cette formation (second et dernier atelier du genre pour la série RGPH 2000) nous fait prévoir de nombreuses candidatures qu'il nous sera nécessaire d'évaluer avec rigueur afin de ne retenir que trente participants, pour conserver à l'atelier un haut niveau technique et un caractère participatif ainsi que pour permettre à chacun d'en tirer un profit maximum.

Nous suggérons donc aux sponsors de soumettre rapidement deux ou trois candidatures détaillées par FAX à l'Equipe d'Appui Technique du FNUAP de Dakar (221- 8.22.83.82) ou au Bureau de Recensement des Etats Unis (1-301-457.30.33 ou rebecca.a.sauer@ccmail.census.gov [e-mail]) qui coordonneront la sélection des participants. La date de clôture des inscriptions est fixée au 30 juillet 1999.

Dissemination Made Easy

Brazil Makes Portuguese Version of Map Viewer
The Instituto Brasileiro de Geografia e Estatística (IBGE) is planning to use the IMPS Map Viewer as a data dissemination tool for their upcoming 2000 Census of Population and Housing. They will prepare sets of summary census data, which they will distribute to the public with the IMPS Map Viewer. Ari Silva, an old friend of IPC from IBGE, worked closely with the IMPS development team to prepare a Portuguese version of Map Viewer.

Philippines Will Disseminate Map Viewer and Cross Tab with Census Data
The National Statistics Office (NSO) of the Philippines is using the IMPS Map Viewer and Cross Tab as data dissemination tools for their 1995 Population Census. The NSO will distribute CD-ROMs with Public Use Files (microdata), summary data, and the IMPS modules. These state-of-the-art data dissemination products were developed by NSO staff working under Val Abuan and Pidio Nogales.